

SEQ SEARCH SUMMARY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 17:42:21 ; Search time 9796 Seconds
(without alignments)
11605.633 Million cell updates/sec

Title: US-09-623-596-1_COPY_1_2623
Perfect score: 2623
Sequence: 1 aagcttcggggaggggcgcg.....tgatgacctcgatttggtg 2623

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

28: em_un:*
 29: em_vi:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_pln:*
 35: em_htg_rod:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. Score | | Query | | DB | ID | Description |
|--------|-----------|--------|-------|--------|----|-----------|--------------------|
| | | | Match | Length | | | |
| | 1 | 2612 | 99.6 | 5348 | 1 | AB041009 | AB041009 Corynebac |
| c | 2 | 2511.6 | 95.8 | 337200 | 1 | AP005280 | AP005280 Corynebac |
| c | 3 | 2511.6 | 95.8 | 349980 | 6 | AX127150 | AX127150 Sequence |
| | 4 | 2016 | 76.9 | 2242 | 6 | AX771905 | AX771905 Sequence |
| | 5 | 1758.6 | 67.0 | 1953 | 6 | AX122456 | AX122456 Sequence |
| | 6 | 1758.6 | 67.0 | 1953 | 6 | BD164573 | BD164573 Novel pol |
| c | 7 | 1208.6 | 46.1 | 306650 | 1 | AP005221 | AP005221 Corynebac |
| c | 8 | 604.4 | 23.0 | 348408 | 1 | BX248358 | BX248358 Corynebac |
| | 9 | 595 | 22.7 | 795 | 6 | AX122457 | AX122457 Sequence |
| | 10 | 595 | 22.7 | 795 | 6 | BD164574 | BD164574 Novel pol |
| c | 11 | 399 | 15.2 | 38859 | 1 | MLCB268 | AL022602 Mycobacte |
| | 12 | 399 | 15.2 | 348450 | 1 | MLEPRTN4 | AL583920 Mycobacte |
| c | 13 | 390 | 14.9 | 15776 | 1 | AE007069 | AE007069 Mycobacte |
| c | 14 | 390 | 14.9 | 37586 | 6 | AX191745 | AX191745 Sequence |
| c | 15 | 390 | 14.9 | 306050 | 1 | BX248341 | BX248341 Mycobacte |
| c | 16 | 390 | 14.9 | 348247 | 15 | BX842579 | Bx842579 Mycobacte |
| | 17 | 155 | 5.9 | 11803 | 1 | AE014760 | AE014760 Bifidobac |
| | 18 | 155 | 5.9 | 349980 | 6 | AX492786 | AX492786 Sequence |
| | 19 | 155 | 5.9 | 349980 | 6 | AX553953 | AX553953 Sequence |
| | 20 | 116.2 | 4.4 | 2951 | 1 | AF123319 | AF123319 Streptomy |
| c | 21 | 116.2 | 4.4 | 321250 | 1 | SC0939111 | AL939111 Streptomy |
| | 22 | 101.6 | 3.9 | 299925 | 1 | AP005045 | AP005045 Streptomy |
| c | 23 | 92.2 | 3.5 | 301205 | 1 | AE017218 | AE017218 Geobacter |
| | 24 | 89.6 | 3.4 | 249150 | 1 | AP005332 | AP005332 Vibrio vu |
| c | 25 | 89.6 | 3.4 | 301442 | 1 | AE016798 | AE016798 Vibrio vu |
| c | 26 | 84 | 3.2 | 208524 | 1 | AE016925 | AE016925 Chromobac |
| c | 27 | 83.8 | 3.2 | 10423 | 1 | AE007547 | AE007547 Clostridi |
| | 28 | 82.6 | 3.1 | 301995 | 1 | AE016779 | AE016779 Pseudomon |
| | 29 | 81.8 | 3.1 | 7983 | 1 | AE013142 | AE013142 Thermoana |
| c | 30 | 81 | 3.1 | 11975 | 1 | AE013120 | AE013120 Thermoana |
| | 31 | 79.8 | 3.0 | 2154 | 1 | AY142812 | AY142812 Heliobaci |
| | 32 | 78.8 | 3.0 | 192000 | 1 | AE001438 | AE001438 Clostridi |
| c | 33 | 78.2 | 3.0 | 300950 | 1 | AP001516 | AP001516 Bacillus |

called fts I

ATCC 13032 genome

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 17:41:51 ; Search time 949 Seconds
(without alignments)
11741.858 Million cell updates/sec

Title: US-09-623-596-1_COPY_1_2623
Perfect score: 2623
Sequence: 1 aagcttcggggaggggcgcg.....tgatgatcctcgatttggtg 2623

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | % Query Match | Length | DB | ID | Description |
|--------|-----|--------|---------------------|--------|----|----------|--------------------|
| | 1 | 2623 | 100.0 | 5347 | 2 | AAZ09633 | Aaz09633 B. lactof |
| c | 2 | 2511.6 | 95.8 | 349980 | 5 | AAH68531 | Aah68531 C glutami |
| | 3 | 2016 | 76.9 | 2242 | 9 | ADD13716 | Add13716 C. glutam |
| | 4 | 1758.6 | 67.0 | 1953 | 5 | AAH67337 | Aah67337 C glutami |
| | 5 | 596.6 | 22.7 | 1827 | 7 | ACA29976 | Aca29976 Prokaryot |
| | 6 | 595 | 22.7 | 795 | 5 | AAH67338 | Aah67338 C glutami |
| | 7 | 403.2 | 15.4 | 1989 | 7 | ACA38015 | Aca38015 Prokaryot |

APPL.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 17:45:31 ; Search time 5965 Seconds
(without alignments)
13131.349 Million cell updates/sec

Title: US-09-623-596-1_COPY_1_2623
Perfect score: 2623
Sequence: 1 aagcttcggggaggggcgcg.....tgatgatcctcgatttggtg 2623

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*

28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | | | % | | | | DB | ID | Description |
|--------|-----|-------|-------|-------|--------|----------|----|----------|-------------|
| | No. | Score | Query | Match | Length | | | | |
| | 1 | 98.6 | 3.8 | 507 | 28 | BH740525 | | BH740525 | cpbav0010 |
| c | 2 | 85.8 | 3.3 | 549 | 28 | BH740502 | | BH740502 | cpbav0008 |
| | 3 | 58.4 | 2.2 | 1133 | 28 | BZ565848 | | BZ565848 | pacs2-164 |
| | 4 | 57.4 | 2.2 | 1297 | 28 | BZ565839 | | BZ565839 | pacs2-164 |
| | 5 | 56.8 | 2.2 | 1138 | 28 | BZ568217 | | BZ568217 | pacs2-164 |
| | 6 | 54.2 | 2.1 | 713 | 14 | CF887190 | | CF887190 | UI-CF-FN0 |
| | 7 | 54 | 2.1 | 1344 | 28 | BH770683 | | BH770683 | LLMGtag43 |
| | 8 | 49.4 | 1.9 | 307 | 10 | BF065979 | | BF065979 | HV_CEB001 |
| c | 9 | 48.6 | 1.9 | 962 | 28 | BZ554173 | | BZ554173 | pacs1-60 |
| | 10 | 44.8 | 1.7 | 1201 | 13 | BX381961 | | BX381961 | BX381961 |
| c | 11 | 44.6 | 1.7 | 1201 | 13 | BX356664 | | BX356664 | BX356664 |
| c | 12 | 42.2 | 1.6 | 1201 | 13 | BX381961 | | BX381961 | BX381961 |
| c | 13 | 42.2 | 1.6 | 1201 | 13 | BX440707 | | BX440707 | BX440707 |
| | 14 | 41.6 | 1.6 | 819 | 28 | CC114410 | | CC114410 | NDL.70I5. |
| c | 15 | 41.6 | 1.6 | 910 | 29 | CNS0060N | | AL065629 | Drosophil |
| | 16 | 41.4 | 1.6 | 600 | 10 | BF006977 | | BF006977 | 1163138 A |
| c | 17 | 41.4 | 1.6 | 609 | 28 | CC063491 | | CC063491 | fgma001d0 |
| | 18 | 41 | 1.6 | 747 | 12 | BI890692 | | BI890692 | ZF637-2-0 |
| c | 19 | 41 | 1.6 | 925 | 29 | CNS0091P | | AL053013 | Drosophil |
| | 20 | 40.4 | 1.5 | 1856 | 9 | AL696950 | | AL696950 | AL696950 |
| c | 21 | 40.2 | 1.5 | 712 | 13 | BX416727 | | BX416727 | BX416727 |
| | 22 | 40.2 | 1.5 | 1201 | 13 | BX360624 | | BX360624 | BX360624 |
| c | 23 | 40 | 1.5 | 1101 | 29 | CNS014Y2 | | AL104756 | Drosophil |
| c | 24 | 39.6 | 1.5 | 1071 | 29 | CNS00EMY | | AL069395 | Drosophil |
| | 25 | 39.6 | 1.5 | 1108 | 13 | BX382084 | | BX382084 | BX382084 |
| | 26 | 39.2 | 1.5 | 812 | 29 | CNS0104Q | | AL153067 | Anopheles |
| | 27 | 39.2 | 1.5 | 839 | 29 | CNS004NB | | AL054280 | Drosophil |
| | 28 | 39 | 1.5 | 462 | 9 | AL822466 | | AL822466 | AL822466 |
| | 29 | 39 | 1.5 | 514 | 14 | CA409494 | | CA409494 | 1031_F -P |
| | 30 | 39 | 1.5 | 517 | 28 | AZ342386 | | AZ342386 | 1M0075K24 |
| | 31 | 39 | 1.5 | 798 | 13 | CA103884 | | CA103884 | SCEZHR108 |
| | 32 | 39 | 1.5 | 914 | 11 | CNS09N4Y | | BX066062 | Single re |
| | 33 | 39 | 1.5 | 917 | 11 | CNS09M71 | | BX064841 | Single re |
| c | 34 | 38.8 | 1.5 | 1201 | 13 | BX376097 | | BX376097 | BX376097 |
| | 35 | 38.6 | 1.5 | 234 | 28 | BH643123 | | BH643123 | 1008052B0 |
| c | 36 | 38.4 | 1.5 | 309 | 9 | AI213021 | | AI213021 | y6e10a1.f |
| c | 37 | 38.4 | 1.5 | 316 | 9 | AA784731 | | AA784731 | g2a02a1.f |
| c | 38 | 38.4 | 1.5 | 513 | 14 | CD211465 | | CD211465 | HS1_60_A0 |
| | 39 | 38.4 | 1.5 | 709 | 9 | AJ558868 | | AJ558868 | AJ558868 |
| | 40 | 38.2 | 1.5 | 656 | 14 | CA186387 | | CA186387 | SCSGST311 |
| | 41 | 38.2 | 1.5 | 710 | 14 | CA216650 | | CA216650 | SCCCST3C1 |
| c | 42 | 38 | 1.4 | 480 | 13 | BQ363085 | | BQ363085 | QV0-ST023 |
| | 43 | 38 | 1.4 | 515 | 13 | BX424977 | | BX424977 | BX424977 |
| | 44 | 38 | 1.4 | 711 | 13 | BU023819 | | BU023819 | QHF12P18. |
| | 45 | 38 | 1.4 | 821 | 14 | CF150908 | | CF150908 | AGENCOURT |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 11, 2004, 11:23:25 ; Search time 5372.72 Seconds
(without alignments)
4687.072 Million cell updates/sec

Title: US-09-623-596-2
Perfect score: 2918
Sequence: 1 MVTRIALVIAGVLIIRLGWV.....TRNTGLPSRVLHPLMILDLL 581

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp

Q=/cgn2_1/USPTO_spool_p/US09623596/runat_09092004_094653_11700/app_query.fasta_1
.1486

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09623596 @CGN_1_1_8050 @runat_09092004_094653_11700 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*

```

11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | | % | | Query | | | ID | Description |
|--------|--------|-------|--------|-------|-----------|--|----|---------------------|
| No. | Score | Match | Length | DB | | | | |
| 1 | 2898 | 99.3 | 5348 | 1 | AB041009 | | | AB041009 Corynebact |
| 2 | 2848 | 97.6 | 1953 | 6 | AX122456 | | | AX122456 Sequence |
| 3 | 2848 | 97.6 | 1953 | 6 | BD164573 | | | BD164573 Novel pol |
| 4 | 2848 | 97.6 | 2242 | 6 | AX771905 | | | AX771905 Sequence |
| c 5 | 2848 | 97.6 | 337200 | 1 | AP005280 | | | AP005280 Corynebact |
| c 6 | 2848 | 97.6 | 349980 | 6 | AX127150 | | | AX127150 Sequence |
| c 7 | 2247 | 77.0 | 306650 | 1 | AP005221 | | | AP005221 Corynebact |
| c 8 | 1676 | 57.4 | 348408 | 1 | BX248358 | | | BX248358 Corynebact |
| c 9 | 1206.5 | 41.3 | 37586 | 6 | AX191745 | | | AX191745 Sequence |
| c 10 | 1206.5 | 41.3 | 306050 | 1 | BX248341 | | | BX248341 Mycobacte |
| c 11 | 1206.5 | 41.3 | 348247 | 15 | BX842579 | | | Bx842579 Mycobacte |
| c 12 | 1203 | 41.2 | 38859 | 1 | MLCB268 | | | AL022602 Mycobacte |
| 13 | 1203 | 41.2 | 348450 | 1 | MLEPRTN4 | | | AL583920 Mycobacte |
| c 14 | 1201.5 | 41.2 | 15776 | 1 | AE007069 | | | AE007069 Mycobacte |
| 15 | 626 | 21.5 | 2951 | 1 | AF123319 | | | AF123319 Streptomy |
| c 16 | 626 | 21.5 | 321250 | 1 | SCO939111 | | | AL939111 Streptomy |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 11, 2004, 09:57:30 ; Search time 516.558 Seconds
(without alignments)
4778.171 Million cell updates/sec

Title: US-09-623-596-2
Perfect score: 2918
Sequence: 1 MVTRIALVIAGVLIIRLGWV.....TRNTGLPSRVLHPLMILDLL 581

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp

Q=/cgn2_1/USPTO_spool_p/US09623596/runat_09092004_094652_11690/app_query.fasta_1
.1486

-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09623596_CGN_1_1_886@runat_09092004_094652_11690 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | % Query Match | Length | DB | ID | Description |
|--------|-----|--------|------------------|--------|----|-------------|--------------------|
| | 1 | 2918 | 100.0 | 5347 | 2 | AAZ09633 | Aaz09633 B. lactof |
| | 2 | 2848 | 97.6 | 1953 | 5 | AAH67337 | Aah67337 C glutami |
| | 3 | 2848 | 97.6 | 2242 | 9 | ADD13716 | Add13716 C. glutam |
| c | 4 | 2848 | 97.6 | 349980 | 5 | AAH68531 | Aah68531 C glutami |
| | 5 | 1674 | 57.4 | 1827 | 7 | ACA29976 | Aca29976 Prokaryot |
| | 6 | 1206.5 | 41.3 | 2040 | 4 | AAH52030 | Aah52030 Mycobacte |
| | 7 | 1206.5 | 41.3 | 2040 | 7 | ACA40605 | Aca40605 Prokaryot |
| c | 8 | 1206.5 | 41.3 | 110000 | 4 | AAI99682_24 | Continuation (25 o |
| | 9 | 1203 | 41.2 | 2028 | 7 | ACA39728 | Aca39728 Prokaryot |
| c | 10 | 1201.5 | 41.2 | 110000 | 4 | AAI99683_24 | Continuation (25 o |
| | 11 | 1199 | 41.1 | 1989 | 7 | ACA38015 | Aca38015 Prokaryot |
| | 12 | 1179.5 | 40.4 | 1937 | 7 | ACA38413 | Aca38413 Prokaryot |
| | 13 | 604 | 20.7 | 349980 | 6 | ABQ81848 | Abq81848 Bifidobac |
| | 14 | 580 | 19.9 | 12131 | 4 | AAS59572 | Aas59572 Propionib |
| | 15 | 580 | 19.9 | 12131 | 7 | ACF64501 | Acf64501 Propionib |
| c | 16 | 526.5 | 18.0 | 110000 | 6 | ABQ69245_21 | Continuation (22 o |
| | 17 | 526.5 | 18.0 | 213251 | 6 | ABQ67193 | Abq67193 Listeria |
| | 18 | 520 | 17.8 | 2178 | 7 | ACA28148 | Aca28148 Prokaryot |
| | 19 | 514.5 | 17.6 | 1977 | 7 | ACA29543 | Aca29543 Prokaryot |
| | 20 | 496 | 17.0 | 1740 | 7 | ACA42602 | Aca42602 Prokaryot |
| | 21 | 494.5 | 16.9 | 2256 | 7 | ACA36564 | Aca36564 Prokaryot |
| c | 22 | 494.5 | 16.9 | 110000 | 6 | ABA03041_21 | Continuation (22 o |
| | 23 | 491 | 16.8 | 1764 | 7 | ACA53863 | Aca53863 Prokaryot |
| | 24 | 484.5 | 16.6 | 1818 | 7 | ACA27874 | Aca27874 Prokaryot |
| | 25 | 474 | 16.2 | 1767 | 7 | ACF71678 | Acf71678 Photorhab |
| | 26 | 474 | 16.2 | 110000 | 7 | ACF67367_48 | Continuation (49 o |
| | 27 | 474 | 16.2 | 110000 | 7 | ACF65387_0 | Acf65387 Photorhab |
| | 28 | 472.5 | 16.2 | 1764 | 7 | ACA35526 | Aca35526 Prokaryot |
| | 29 | 465.5 | 16.0 | 1746 | 7 | ACA44262 | Aca44262 Prokaryot |
| | 30 | 465 | 15.9 | 2298 | 8 | ADB11843 | Adb11843 Alloiococ |
| | 31 | 465 | 15.9 | 2298 | 8 | ADB11845 | Adb11845 Alloiococ |
| c | 32 | 465 | 15.9 | 110000 | 8 | ADB12064_13 | Continuation (14 o |
| | 33 | 460 | 15.8 | 1743 | 7 | ACA53408 | Aca53408 Prokaryot |
| | 34 | 459.5 | 15.7 | 1728 | 7 | ACA45661 | Aca45661 Prokaryot |
| | 35 | 457.5 | 15.7 | 2226 | 4 | AAS51352 | Aas51352 Enterococ |
| | 36 | 457.5 | 15.7 | 2229 | 4 | AAS52900 | Aas52900 Enterococ |
| | 37 | 457 | 15.7 | 2226 | 7 | ACA18435 | Aca18435 Prokaryot |
| | 38 | 454 | 15.6 | 1714 | 7 | ACA31790 | Aca31790 Prokaryot |
| | 39 | 453.5 | 15.5 | 1767 | 4 | AAS52291 | Aas52291 E. coli D |
| | 40 | 453.5 | 15.5 | 1767 | 5 | AAH81456 | Aah81456 Escherich |
| | 41 | 453.5 | 15.5 | 1767 | 7 | ACA32355 | Aca32355 Prokaryot |
| | 42 | 453.5 | 15.5 | 1767 | 7 | ACA51127 | Aca51127 Prokaryot |
| | 43 | 449.5 | 15.4 | 1750 | 7 | ACA49176 | Aca49176 Prokaryot |
| | 44 | 449.5 | 15.4 | 1845 | 7 | ACA23990 | Aca23990 Prokaryot |
| | 45 | 444.5 | 15.2 | 1746 | 4 | AAS56356 | Aas56356 Salmonell |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 11, 2004, 14:39:16 ; Search time 3286.03 Seconds
(without alignments)
5279.912 Million cell updates/sec

Title: US-09-623-596-2
Perfect score: 2918
Sequence: 1 MVTRIALVIAGVLIIRLGWV.....TRNTGLPSRVLHPLMILDLL 581

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp

Q=/cgn2_1/USPTO_spool_p/US09623596/runat_09092004_094653_11712/app_query.fasta_1
.1486

-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09623596@CGN_1_1_6628@runat_09092004_094653_11712 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*

```

11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | | | % | | Query | | DB | ID | Description |
|--------|-----|-------|-------|--------|-------|----------|----|----------|-------------|
| | No. | Score | Match | Length | | | | | |
| c | 1 | 342 | 11.7 | 507 | 28 | BH740525 | | BH740525 | cpbav0010 |
| | 2 | 244 | 8.4 | 6303 | 28 | BH771012 | | BH771012 | LLMGtag73 |
| | 3 | 237 | 8.1 | 1344 | 28 | BH770683 | | BH770683 | LLMGtag43 |
| | 4 | 227 | 7.8 | 1133 | 28 | BZ565848 | | BZ565848 | pacs2-164 |
| | 5 | 220 | 7.5 | 1297 | 28 | BZ565839 | | BZ565839 | pacs2-164 |
| | 6 | 219 | 7.5 | 1138 | 28 | BZ568217 | | BZ568217 | pacs2-164 |
| | 7 | 214.5 | 7.4 | 1469 | 28 | BZ572326 | | BZ572326 | msh2_2585 |
| c | 8 | 209 | 7.2 | 962 | 28 | BZ554173 | | BZ554173 | pacs1-60_ |
| | 9 | 204 | 7.0 | 812 | 29 | CNS0104Q | | AL153067 | Anopheles |
| c | 10 | 190 | 6.5 | 785 | 28 | BH795496 | | BH795496 | 30B12LLI0 |
| | 11 | 189.5 | 6.5 | 713 | 14 | CF887190 | | CF887190 | UI-CF-FN0 |
| c | 12 | 178.5 | 6.1 | 930 | 28 | AF036000 | | AF036000 | AF036000 |
| | 13 | 177.5 | 6.1 | 609 | 28 | CC063491 | | CC063491 | fgma001d0 |
| c | 14 | 176 | 6.0 | 1060 | 28 | AF095377 | | AF095377 | AF095377 |
| | 15 | 168 | 5.8 | 958 | 28 | AF029629 | | AF029629 | AF029629 |
| | 16 | 163 | 5.6 | 826 | 28 | BZ571950 | | BZ571950 | msh2_216. |
| | 17 | 154.5 | 5.3 | 307 | 10 | BF065979 | | BF065979 | HV_CEb001 |
| c | 18 | 152 | 5.2 | 1401 | 28 | BZ552366 | | BZ552366 | pacs1-60_ |
| c | 19 | 148 | 5.1 | 528 | 14 | CF324573 | | CF324573 | HDN--06-O |
| c | 20 | 146.5 | 5.0 | 1342 | 28 | BZ553468 | | BZ553468 | pacs1-60_ |
| c | 21 | 137.5 | 4.7 | 1113 | 28 | BZ558395 | | BZ558395 | pacs1-60_ |
| c | 22 | 137.5 | 4.7 | 1600 | 28 | BH770611 | | BH770611 | LLMGtag37 |
| | 23 | 137 | 4.7 | 1030 | 28 | BZ562490 | | BZ562490 | pacs2-164 |
| | 24 | 133.5 | 4.6 | 966 | 12 | BG676279 | | BG676279 | 602622680 |
| | 25 | 133 | 4.6 | 1161 | 28 | BZ567531 | | BZ567531 | pacs2-164 |
| c | 26 | 131 | 4.5 | 2396 | 11 | AK029491 | | AK029491 | Mus muscu |
| | 27 | 129.5 | 4.4 | 1002 | 28 | BZ576640 | | BZ576640 | msh2_5032 |
| | 28 | 127.5 | 4.4 | 3664 | 29 | AY407030 | | AY407030 | Mus muscu |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 17:46:11 ; Search time 204 Seconds
(without alignments)
7135.472 Million cell updates/sec

Title: US-09-623-596-1_COPY_1_2623
Perfect score: 2623
Sequence: 1 aagcttcggggaggggcgcg.....tgatgatcctcgatttg 2623

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | | | % | | Query | | | | ID | Description |
|--------|-----|-------|-------|---------|-------|---------------------|--|--|----|-------------------|
| | No. | Score | Match | Length | DB | | | | | |
| c | 1 | 390 | 14.9 | 4403765 | 3 | US-09-103-840A-2 | | | | Sequence 2, Appli |
| c | 2 | 390 | 14.9 | 4411529 | 3 | US-09-103-840A-1 | | | | Sequence 1, Appli |
| | 3 | 71.6 | 2.7 | 1782 | 4 | US-09-252-991A-8703 | | | | Sequence 8703, Ap |
| | 4 | 71.6 | 2.7 | 1830 | 4 | US-09-252-991A-8916 | | | | Sequence 8916, Ap |
| | 5 | 63.8 | 2.4 | 1815 | 4 | US-09-252-991A-7621 | | | | Sequence 7621, Ap |
| c | 6 | 63.8 | 2.4 | 1932 | 4 | US-09-252-991A-7931 | | | | Sequence 7931, Ap |
| c | 7 | 56.4 | 2.2 | 960 | 4 | US-09-252-991A-9077 | | | | Sequence 9077, Ap |
| | 8 | 51.2 | 2.0 | 2232 | 4 | US-09-134-000C-1279 | | | | Sequence 1279, Ap |
| c | 9 | 50 | 1.9 | 7218 | 1 | US-08-232-463-14 | | | | Sequence 14, Appl |
| | 10 | 45.8 | 1.7 | 1764 | 4 | US-09-489-039A-5156 | | | | Sequence 5156, Ap |
| | 11 | 44 | 1.7 | 1812 | 4 | US-09-543-681A-2803 | | | | Sequence 2803, Ap |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 20:46:57 ; Search time 1188 Seconds
(without alignments)
11109.162 Million cell updates/sec

Title: US-09-623-596-1_COPY_1_2623
Perfect score: 2623
Sequence: 1 aagcttcggggaggggcgcg.....tgatgatcctcgatttggtt 2623

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | | | % | | Query | | | | DB | ID | Description |
|--------|-------|--------|--------|---------|-------|----------------------|--|--|----|----|-------------------|
| | | | | | | | | | | | |
| No. | Score | Match | Length | | | | | | | | |
| c | 1 | 2511.6 | 95.8 | 3309400 | 9 | US-09-738-626-1 | | | | | Sequence 1, Appli |
| | 2 | 1758.6 | 67.0 | 1953 | 9 | US-09-738-626-2372 | | | | | Sequence 2372, Ap |
| | 3 | 596.6 | 22.7 | 1827 | 13 | US-10-282-122A-17846 | | | | | Sequence 17846, A |
| | 4 | 595 | 22.7 | 795 | 9 | US-09-738-626-2373 | | | | | Sequence 2373, Ap |
| | 5 | 403.2 | 15.4 | 1989 | 13 | US-10-282-122A-25885 | | | | | Sequence 25885, A |
| | 6 | 399 | 15.2 | 2028 | 13 | US-10-282-122A-27598 | | | | | Sequence 27598, A |
| | 7 | 390 | 14.9 | 2040 | 9 | US-09-712-363-84 | | | | | Sequence 84, Appl |
| | 8 | 390 | 14.9 | 2040 | 13 | US-10-282-122A-28475 | | | | | Sequence 28475, A |
| | 9 | 379.6 | 14.5 | 1937 | 13 | US-10-282-122A-26283 | | | | | Sequence 26283, A |
| | 10 | 155 | 5.9 | 2256646 | 17 | US-10-470-565-1 | | | | | Sequence 1, Appli |
| | 11 | 101.6 | 3.9 | 1959 | 15 | US-10-156-761-6093 | | | | | Sequence 6093, Ap |
| | 12 | 101.6 | 3.9 | 9025608 | 15 | US-10-156-761-1 | | | | | Sequence 1, Appli |
| | 13 | 84.8 | 3.2 | 2178 | 13 | US-10-282-122A-16018 | | | | | Sequence 16018, A |
| | 14 | 83.8 | 3.2 | 1818 | 13 | US-10-282-122A-15744 | | | | | Sequence 15744, A |
| | 15 | 82.6 | 3.1 | 1746 | 13 | US-10-282-122A-32132 | | | | | Sequence 32132, A |
| | 16 | 73.6 | 2.8 | 2731748 | 17 | US-10-297-465A-1 | | | | | Sequence 1, Appli |
| | 17 | 72.2 | 2.8 | 1842 | 13 | US-10-282-122A-14100 | | | | | Sequence 14100, A |
| | 18 | 71.8 | 2.7 | 1919 | 9 | US-09-974-300-1666 | | | | | Sequence 1666, Ap |
| | 19 | 70 | 2.7 | 1698 | 9 | US-09-815-242-7770 | | | | | Sequence 7770, Ap |
| | 20 | 70 | 2.7 | 1698 | 13 | US-10-282-122A-30183 | | | | | Sequence 30183, A |
| | 21 | 69.2 | 2.6 | 1845 | 13 | US-10-282-122A-11860 | | | | | Sequence 11860, A |
| | 22 | 67.4 | 2.6 | 2190 | 13 | US-10-282-122A-15800 | | | | | Sequence 15800, A |
| | 23 | 64 | 2.4 | 1863 | 13 | US-10-282-122A-12918 | | | | | Sequence 12918, A |
| | 24 | 63.8 | 2.4 | 1740 | 13 | US-10-282-122A-30472 | | | | | Sequence 30472, A |
| | 25 | 62.8 | 2.4 | 1716 | 13 | US-10-282-122A-29337 | | | | | Sequence 29337, A |
| | 26 | 59.2 | 2.3 | 1746 | 13 | US-10-282-122A-29883 | | | | | Sequence 29883, A |
| | 27 | 58.4 | 2.2 | 1719 | 13 | US-10-282-122A-15195 | | | | | Sequence 15195, A |
| | 28 | 56.6 | 2.2 | 1728 | 13 | US-10-282-122A-33531 | | | | | Sequence 33531, A |
| | 29 | 53.2 | 2.0 | 1959 | 13 | US-10-282-122A-40409 | | | | | Sequence 40409, A |
| | 30 | 51.6 | 2.0 | 1767 | 13 | US-10-282-122A-38997 | | | | | Sequence 38997, A |
| | 31 | 51.4 | 2.0 | 2247 | 13 | US-10-282-122A-36335 | | | | | Sequence 36335, A |
| | 32 | 51.2 | 2.0 | 2226 | 9 | US-09-815-242-3934 | | | | | Sequence 3934, Ap |
| | 33 | 51.2 | 2.0 | 2226 | 13 | US-10-282-122A-6305 | | | | | Sequence 6305, Ap |
| | 34 | 51.2 | 2.0 | 2229 | 9 | US-09-815-242-6537 | | | | | Sequence 6537, Ap |
| | 35 | 51 | 1.9 | 1764 | 13 | US-10-282-122A-41733 | | | | | Sequence 41733, A |
| | 36 | 49.8 | 1.9 | 1750 | 13 | US-10-282-122A-37046 | | | | | Sequence 37046, A |
| | 37 | 49.4 | 1.9 | 1767 | 9 | US-09-741-669-255 | | | | | Sequence 255, App |
| | 38 | 49.4 | 1.9 | 1767 | 9 | US-09-815-242-5928 | | | | | Sequence 5928, Ap |
| | 39 | 49.4 | 1.9 | 1767 | 13 | US-10-282-122A-20225 | | | | | Sequence 20225, A |
| | 40 | 48.4 | 1.8 | 1714 | 13 | US-10-282-122A-19660 | | | | | Sequence 19660, A |
| | 41 | 47.2 | 1.8 | 2256 | 13 | US-10-282-122A-38496 | | | | | Sequence 38496, A |
| | 42 | 46.4 | 1.8 | 2256 | 13 | US-10-282-122A-24434 | | | | | Sequence 24434, A |
| | 43 | 46 | 1.8 | 1743 | 13 | US-10-282-122A-41278 | | | | | Sequence 41278, A |
| | 44 | 45.8 | 1.7 | 1764 | 13 | US-10-282-122A-23396 | | | | | Sequence 23396, A |
| | 45 | 44.8 | 1.7 | 1068 | 9 | US-09-974-300-6191 | | | | | Sequence 6191, Ap |

ALIGNMENTS

RESULT 1
US-09-738-626-1/c